IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | | |
|---|---|---|---|
| Appellant: | Chrissan et al. | Examiner: | Azad, A. |
| Serial No.: | 09/392,124 | Group Art Unit: | 2654 |
| Filed: | September 8, 1999 | Docket No.: | 8X8S.239PA |
| Title: | Varying Pulse Amplitude Multi-Pulse Analysis Speech Processor And Method | | |

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence and the papers, as described hereinabove, are being deposited in the United States Postal Service **in triplicate,** as first class mail, in an envelope addressed to: Mail Stop Appeal Brief – Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on September 7, 2004.
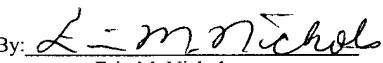
By: _Erin M. Nichols_
Erin M. Nichols

## APPEAL BRIEF

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This is an Appeal Brief submitted pursuant to 37 C.F.R. § 1.192 for the above-referenced patent application. Please charge Deposit Account No. 50-0996 (8X8S.239PA) in the amount of $330 for this brief in support of appeal as indicated in 37 C.F.R. § 1.17(c). If necessary, authority is given to charge/credit deposit account 50-0996 (8X8S.239PA) any additional fees/overages in support of this filing.

## I. Real Party in Interest

The real party in interest is 8x8, Inc., formerly Netergy Microelectronics, Inc., having a principal place of business at 2445 Mission College Boulevard, Santa Clara, CA 95054. The above-referenced patent application is assigned to 8x8, Inc.

## II. Related Appeals and Interferences

Appellant is unaware of any related appeals or interferences.

## III. Status of Claims

Claims 1-32 are presented for appeal. Claims 1-27 and 29-32 stand rejected under 35 U.S.C. § 103(a) over *Bialik et al.* (U.S. Pat. No. 5,568,588) in view of *Adoul et al.* (U.S. Patent No. 5,754,976); and claim 28 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over *Bialik* in view of *Adoul* and further in view of *Sklar* (Digital Communications Fundamentals and Application). The pending claims under appeal, as presently amended, may be found in the attached Appendix of Appealed Claims.

## IV. Status of Amendments

The application was originally filed on September 8, 1999, including 32 claims. A first Office Action was mailed on January 30, 2002, and in reply, an Office Action Response was filed on July 30, 2002. A final Office Action was mailed on October 22, 2002, and in reply, a Response To Final Office Action and a Notice of Appeal were concurrently filed by facsimile on January 16, 2003. An Advisory Action, which included a new citation (U.S. Patent No. 3,624,302) was mailed on February 20, 2003. On March 17, 2003, an Appeal Brief was filed. An Examiner's Answer was mailed on June 17, 2003, and in reply, a Reply Brief was filed on August 18, 2003. An Office Action was mailed on November 24, 2003, reopening prosecution, and in reply, an Office Action Response was filed on January 14, 2004. A final Office Action was mailed on April 7, 2004, and in reply, an Office Action Response After Final was filed on June 7, 2004. A Notice of Appeal was filed on July 7, 2004, and an Advisory Action was mailed on July 12, 2004.

## V. Summary of Invention

The present invention is directed to a speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector. The method includes generating a plurality of sequences of variable-amplitude pulses from the target vector and the short term characteristics, where each of the sequences has a different average amplitude value, and outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

The present invention is directed to providing a significant improvement on the multi-pulse speech analysis and synthesis ("MPA") teachings of the *Bialik* '588 reference.

Multi-pulse speech analysis and synthesis typically involves dividing the incoming speech signals into frames and then analyzing each frame to determine its representative components, for example, using a frame analyzer to determine the short-term and long-term characteristics of the speech signal. Typically, one, both, or neither of the long- and short-term predictor contributions are subtracted from the input frame, leaving a target vector whose shape has to be characterized from a multiplicity of samples.

As discussed in the background section of Appellant's Specification, one particular MPA approach is described by the '588 reference. The target vector is modeled by a plurality of pulses of equal amplitude, varying location and varying sign (positive and negative). To select each equal-amplitude pulse, a pulse is placed at each sample location and the effect of the pulse, defined by passing the pulse through a filter defined by the LPC coefficients, is determined. The pulse which provides the filter output that most closely matches the target vector is selected and its effect is removed from the target vector, thereby generating a new target vector. The process continues until a predetermined number of pulses have been found. For storage or transmission purposes, the result of the MPA analysis is a collection of pulse locations, pulse signs (positive or negative), and a quantized value of the equal pulse amplitude in each sequence.

In the prior art, the MPA output typically specifies the resulting pulse locations, but not the order in which they were chosen. It also specifies only one gain parameter, so the decoder must reconstruct the pulse sequence using equal amplitudes for all the pulses in the sequence.

According to an example embodiment, the present invention significantly improves over the '588 reference by performing a more accurate MPA analysis; an MPA analysis that, from a maximum-likelihood standpoint, has a much better opportunity for determining the best possible pulse sequence to match the target vector because the pulse sequence is reconstructed using varying amplitude pulses in the sequence. By determining a better match to the target, the perceptual quality of the reconstructed speech is significantly improved.

## VI. Issues for Review

Issue I:  Are the § 103(a) rejections of claims 1-32 proper when the Examiner failed to present a *prima facie* case of obviousness?

Issue II:  Is the § 103(a) rejection of claims 1-27 and 29-32 proper when the Examiner's proposed modification of the '588 reference would frustrate the purpose and operation of the '588 reference?

Issue III:  Is the § 103(a) rejection of claim 28 proper when the Examiner fails to present a *prima facie* rejection by failing to present a combination of references that corresponds to the claimed invention and failing to present evidence of motivation for the modification proposed by the Examiner?

Issue IV:  Are the § 103(a) rejections of claims 1-32 proper when the Examiner failed to take note of Appellant's arguments presented in the Office Action Response filed on January 14, 2004, and answer the substance thereof, as required by MPEP § 707.07(f)?

## VII. Grouping of Claims

The claims as now presented do not stand and fall together and are separately patentable for the reasons discussed in the Argument.  For purposes of this appeal, the claims should be grouped as follows: Group I - claims 1-27 and 29-32; and Group II – claim 28.

## VIII. Argument

Appellant submits that the claims of groups I – II are patentably distinguishable from each other and from the cited prior art references.  The claims in group I are patentable over the prior art, because they include subject matter that is not taught or suggested by any of the references cited, including generating from a target vector and short term characteristics, a plurality of sequences of variable-amplitude pulses.  The claim of group II is separately patentable over the other claim group because it is directed to subject matter that includes a pulse-train sequence modification function based upon the exponential function, which is not necessarily present in the other claim groups and not taught by the cited prior art.

4

**Issue I:** **The § 103(a) rejections of claims 1-32 are not proper when the Examiner failed to present a *prima facie* case of obviousness.**

A. The § 103(a) Rejections Are Supported By An Unexplained And Illogical Rationale

The Examiner's Section 103(a) rejections fail to present a *prima facie* case of obviousness because Appellant's claims cannot be read to cover the '588 reference, with or without the Examiner's proposed modifications in view of the '976 *Adoul* reference. With reference to Fig. 1 of the '588 reference, the Examiner alleges that the claim (*e.g.*, claim 1) limitations "*generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses*," reads on the '588 reference's elements 10, 13, 20 and 38. Perhaps in view of the clear teaching in each of the '588 embodiments and the '588 claims (*see also* equal-amplitude pulses in Figs. 3A, 3B, 4A and 4B), the Examiner has correctly admitted that none of these cited elements (or any other aspects of the '588 reference) teach "*sequences of variable-amplitude pulses*" (*see* Final Office Action, page 3, last paragraph).

In a hindsight attempt to overcome this deficiency, the Examiner has proposed modifying these cited elements of the '588 reference in view of the alleged teaching of codevector-waveform pulse positions from the '976 reference. Failing to support the rejection, the Examiner has not, however, explained how this could be accomplished. Moreover, deduction would suggest that this proposed modification is somehow achieved without modifying elements 10 or 13. This logic follows since elements 10 and 13 of the '588 reference provide the short-term characteristics (from element 10) and the target vector (from element 13) from which "*a plurality of sequences of variable-amplitude pulses*" are to be generated. Further, element 20 is described by the '588 reference as merely to "determine[s] the sample location of a first pulse in accordance with [known MPA] multi-pulse analysis techniques" ('588 reference, Col. 4, lines 10-11); as such, it seems untenable that the Examiner would be suggesting that the skilled artisan would be led to change element 20 so that, instead of determining the sample locations of the pulses, it would be modified to generate "*sequences of variable-amplitude pulses.*" Accordingly, the Examiner's proposed modification would have to be achieved by modifying element 38 with the alleged '976 teachings of codevector-waveform pulse positions.

In an abundance of caution, Appellant assumes that the Examiner intended for element 38 to mean element 28, since element 38 is merely an output line that carries the overall result of the analysis (*i.e.*, the pulse sequence that matches the target vector). With this assumption, the

5

Examiner's proposed modification would have to be achieved by modifying element 28. Element 28 is the "target vector matcher" that generates the above-mentioned overall result. Thus, Appellant assumes that the Examiner is basing the rejection on the proposed modification of the "target vector matcher 28" with certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference.

Appellant respectfully submits that this proposed modification could not result, even with lengthy research and explanations, in "target vector matcher 28" operating to generate "*sequences of variable-amplitude pulses.*" The proposed modification to the "target vector matcher 28" would have to somehow operate as a function of an input signal that changes the amplitude of the pulses in each given sequence. However, with this proposed modification, the "target vector matcher 28" still acts based on the following two inputs: the target vector from element 13, and each pulse sequence provided by line 34. As is typical for every such standard MPA implementation (Col. 1, lines 35-45), the target vector is provided as an input solely as a reference against which the match (estimation) is made, and each pulse sequence provided by line 34 is a sequence of equal amplitude pulses as illustrated, *e.g.*, in Figs. 3A, 3B, 4A, and 4B, the discussions at Col. 2, lines 50-51, Col. 6, lines 8-29, and each issued claim of the '588 reference.

Accordingly, in view of the rationale provided by the Examiner or by deduction, Appellant respectfully submits that this proposed modification does not result in a hypothetical embodiment which corresponds to Appellant's claimed invention including, for example, "*generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses.*"

### B. The § 103(a) Rationale Would Combine Competing And Incompatible Systems

The Examiner's rejections are based on the MPA type of speech coding approach (as exemplified by the '588 reference) but as somehow modified by certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference. Codevector-waveform pulse positions are used in another type of speech coding system known as the Code-Excited Linear Predictive ("CELP") coding system. Each of the embodiments described by the '976 reference uses and implements the CELP type of coding system as is clearly supported by its "Background of the Invention" section and its detailed description of the algebraic codebook implementations. Indeed, the algebraic codebook implementations correspond to one of the two types of known CELP coding

systems described in the "Background of the Invention" section ('976 reference at Col. 2, line 5 *et seq.*).

Appellant respectfully submits that the relied-upon aspects of this '976 CELP coding system are incompatible with the MPA speech coding approach taught by the '588 reference. The differences between these systems are notoriously well known to the skilled artisan, and the Examiner has failed to cite any evidence or even explain what certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference are being used to modify the MPA speech coding approach taught by the '588 reference. Accordingly, these cited references are directed to incompatible speech encoding methods. *See* attached article "Hybrid Codecs".

At page 3 of the final Office Action, the Examiner asserts that the '976 reference teaches an algebraic codebook search method that involves "sequences of variable-amplitude pulses" by citing the encoding principle of the '976 reference. These arguments by the Examiner further show this incompatibility. The Examiner agrees that the purpose of the '976 teachings is to pre-establish a function $S_p$. This is accomplished via the algorithm outlined at Col. 12, line 34 – Col. 14, line 26, which is used to achieve "restraining the subset of codevectors $A_k$ being searched" in the codebook. *See* Col. 14, lines 19-26. However, the '588 reference does not have any such codebook to search. The '588 reference uses target vector matcher 28 on-the-fly and never searches (or mentions) any codebook; this follows as the '588 and '976 methods are entirely different and incompatible aspects of speech coding systems. *See* attached article "Hybrid Codecs" page 2, last paragraph, to page 3, first full paragraph. Further, the Examiner fails to correlate these disparate methods of speech coding. A traditional code-excited linear predictive (CELP) system is shown in the attached block diagram. *See* "CELP Coding of Speech," page 2. The Examiner has failed to identify how the traditional codebooks would be utilized by the '588 method in the proposed combination. Thus, the Examiner has failed to present correspondence to the claimed plurality of sequences of variable-amplitude pulses and failed to present a combination of references that is directed to the same method of speech encoding. Without either of these requirements being met, the Examiner has failed to present a *prima facie* rejection and the Section 103(a) rejections cannot be maintained. Appellant respectfully requests that the rejections be reversed.

> **Issue II:** **The § 103(a) rejection of claims 1-27 and 29-32 is not proper when the Examiner's proposed modification of the '588 reference would frustrate the purpose and operation of the '588 reference.**

7

Similar to the previous appeal of these claims, a primary issue is the '588 reference's requirement that each sequence of pulses (or train of pulses) have the same amplitude. Evidence of the significance of this aspect of the '588 reference's teaching, may be seen in that each of the three '588 embodiments as well as each '588 claim includes this requirement. Each of the independent claims is directed to "a plurality of sequences of equal amplitude" (claims 1 and 9), "a sequence of equal amplitudes" (claims 2 and 12), "sign trains of equal amplitude" (claim 5), "variable sign trains of equal amplitude" (claim 10), "trains having the same amplitude level" (claim 15), and "pulses having the same amplitude" (claim 16). With respect to the '588 embodiments, FIGs. 3A, 3B, 4A and 4B, illustrate the equal-amplitude pulse trains and operation of the first embodiment. *See* Col. 2, lines 43-54 (brief description).

In maintaining the prior art rejection, the Examiner improperly attempts to overcome deficiencies in the '588 reference. The Examiner's rejection proposes using the '976 reference's teachings regarding variable amplitude pulses in the '588 reference's processing system.

Appellant has repeatedly shown that the proposed combination of alleged prior art is improper because it would frustrate the purpose and operation of the '588 reference. *See* MPEP § 2143.01 (when a proposed modification would render the teachings being modified unsatisfactory for their intended purpose, then there is no suggestion or motivation to make the proposed modification under 35 U.S.C. § 103(a)).

As with the previous appeal of these claims, the Examiner refuses to accept that the '588 reference requires each sequence of pulses (or train of pulses) to have the same amplitude, as discussed above. The Examiner has not, in any Office Action, explained how the '976 teachings would be combined with the '588 embodiment, thereby failing to comply with 35 U.S.C. § 132 and further, precluding Appellant from considering and responding to the merits of the proposed combination. Notwithstanding this lack of compliance with 35 U.S.C. § 132, Appellant surmised that the proposed modification was to replace the '588 processing of the plurality of single gain pulses with the '976 reference's pulse encoding principle. If the Examiner is suggesting that the differing gain values should be replaced by the '976's amplitude selector 112, which provides a specific function $S_p$ (*see* '976 reference at Col. 12, lines 24-33), the '588's purpose (matching the target vector via performing single gain multi-pulse analysis a number of times) would be destroyed. If the amplitude selector 112 replaces the multiple single gain multi-pulse analysis,

then the gain is "pre-established" per the '976 teachings; therefore, the gain is identified without a recurring process (see '588 reference Col. 1, lines 49-55 and Col. 2, lines 1-6). This '976 encoding principal would not function at all in the '588 embodiment and, adopting the Office Action's interpretation of the '976 variable-gain encoding principal, the '588 embodiment would no longer have the required single-level pulse sequences. In this regard, the proposed combination would frustrate the operation and purpose of the '588 embodiment. Thus, the Office Action's proposed combination is improper and the rejection cannot be maintained.

Moreover, the Examiner's alleged motivation for the proposed combination is illogical and untenable. The Examiner provides the unsupported statement that the skilled artisan would combine the cited references to obtain a "very good performance" "without paying a heavy price." The Examiner's failure to explain how a "very good performance" would be achieved should be clear in view of the above-discussed resulting inoperable device constructed via this hindsight rejection. The '588 reference teaches an embodiment that attempts to match a target vector by performing single gain multi-pulse analysis a number of times, each with a different gain level. The '976 approach, upon which the Examiner is relying, is an encoding technique that uses a special amplitude selector 112 (Figs. 3A, B and C) to provide a pre-established function (i.e., a pre-established gain) for a pre-assigned relationship to the speech signal (see Col. 12, lines 29-33). Replacing the '588 multi-pulse analysis approach (using different gain levels) with the '976 pre-established function would eliminate the recurring process for target vector matching and destroy the '588 method. The Office Action fails to present any evidence of the alleged motivation and the cited teachings would certainly not be motivated. Without a presentation of evidence of motivation, the Section 103(a) rejections cannot be maintained.

**Issue III:** **The § 103(a) rejection of claim 28 is not proper when the Examiner fails to present a _prima facie_ rejection by failing to present a combination of references that corresponds to the claimed invention and failing to present evidence of motivation for the modification proposed by the Examiner.**

The Examiner failed to present a combination of references that correspond to the claimed invention and failed to present evidence of motivation for the proposed modification of the '588 reference. The Examiner further fails to comply with 35 U.S.C. § 132 because no explanation has been given as to how the teachings of _Sklar_ would be combined with the above-

9

discussed modified '588 embodiment, and because the Examiner fails to cite any evidence in support of the notion that the skilled artisan would be led by the prior art to implement this asserted combination of teachings. In this regard, the rejection has not afforded Appellant an opportunity to consider and respond to the merits of this proposed combination of three different teachings. *See* 35 U.S.C. § 132. Moreover, the proposed modification (to replace the '588 processing of the plurality of single gain pulses with the '976 reference's pulse encoding principle and also the cited teaching of *Sklar*) would neither correspond to Appellant's claimed invention (as explained above), would frustrate the purpose and teachings of the '588 reference, and would not (contrary to the unexplained assertion in the Office Action) necessarily result in improved output speech quality. Appellant fails to recognize any evidence that has been presented by the Examiner that such a combination of prior art teachings has ever been suggested or even considered.

The Examiner erroneously asserts that the skilled artisan would be lead by the prior art to modify the '588 reference so that it uses an exponential modification function to provide pulses of varying amplitude in each pulse-train sequence because this would allegedly improve output speech quality. Appellant submits that modifying the '588 reference in this regard would not improve output speech quality because the functional blocks described by the '588 reference would still operate under the design principle that the pulses in each pulse-train sequence have the same amplitude. Thus, the Examiner's assertion is illogical.

The Examiner's assertion in this regard would also undermine the operation and objectives of the '588 reference. As stated in the Summary of the '588 reference and discussed above, each pulse sequence has a single gain level and each pulse sequence is processed as this "single gain pulse sequence" (Col. 2, line 11). The Examiner's proposed modification, however, would result in a different set of objectives, in an inaccurate "perceptual weighting filter" (Col. 2, lines 11-12), an inoperable gain selector, and due to a set of unmappable gain levels for each pulse sequence, such pulse sequences which would not be identifiable to "minimize the energy of the error vector and its corresponding gain level" (Col. 2, lines 12-15). Such a destructive combination is improper and fails to indicate the requisite motivation to support a Section 103(a) rejection. The Examiner has not presented a *prima facie* case of rejection; therefore, Appellant submits that the rejection should be reversed.

10

**Issue IV:** The § 103(a) rejections of claims 1-32 are not proper when the Examiner failed to take note of Appellant's arguments presented in the Office Action Response filed on January 14, 2004, and answer the substance thereof, as required by MPEP § 707.07(f).

The Examiner's April 7th Office Action failed to address Appellant's arguments, while repeating the previous rejections. MPEP § 707.07(f) states, in pertinent part, the following:

> Where the requirements are traversed, or suspension thereof requested, the examiner should take proper reference thereto in his or her action on the amendment. Where the applicant traverses any rejection, the examiner should, if he or she repeats the rejection, take note of the applicant's argument and answer the substance of it. If a rejection of record is to be applied to a new or amended claim, specific identification of that ground of rejection, as by citation of the paragraph in the former Office letter in which the rejection was originally stated, should be given.

In this regard, MPEP § 707.07(f) indicates that the Examiner should take note of Appellant's arguments regarding the impropriety of the proposed combination and answer the substance of it. This is consistent with the purpose of aiding the Appellant in judging the propriety of continuing the prosecution, as indicated in 37 C.F.R. § 1.104(a)(2).

In this instance, the Examiner did not comply with this requirement, and Appellant was not afforded the opportunity to judge the propriety of the Section 103(a) rejections and to form a response thereto. For example, the Examiner failed to respond to Appellant's arguments regarding the proposed combination's lack of motivation due to the resulting frustration of the '588 reference's purpose and operation at pages 4-5 in Appellant's January 14th Office Action Response and at pages 2-3 of Appellant's June 7th Office Action Response After Final. Therefore, Appellant requests that the finality of the Office Action mailed on April 7, 2004, be removed, that the Examiner take reference to the Appellant's arguments, and that the Appellant have an opportunity to respond thereto, should the rejection be maintained.

11

## IX. Conclusion

In view of the above, Appellant submits that the rejections are improper, the claimed invention is patentable, and that the rejections of claims 1-32 should be reversed. Appellant respectfully requests reversal of the rejections as applied to the appealed claims and allowance of the entire application.

Authority to charge the undersigned's deposit account was provided on the first page of this brief.

Respectfully submitted,

CRAWFORD MAUNU PLLC
1270 Northland Drive – Suite 390
St. Paul, MN 55120
(651) 686-6633

By: _____
Name:  Robert J. Crawford
Reg. No. 32,122

# APPENDIX OF APPEALED CLAIMS (S/N 09/392,124)

1.     In a speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, a method of analyzing the input speech signal comprising:

generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

2.     A system according to claim 1, wherein the target vector is matched using a perceptual weighting criterion.

3.     A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

means for generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

means for outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

4.     A system according to claim 3, wherein the target vector is matched using a perceptual weighting criterion.

5.     A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

an analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude pulses, each of said sequences having a different average amplitude value;

the analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

6.      A system according to claim 5, wherein the target vector is matched using a perceptual weighting criterion.

7.      A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude, variable-sign and variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs;

the multi-pulse analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude, variable-sign, variably-spaced pulses which, according to a maximum likelihood criterion, most closely represents the target vector.

8.      A system according to claim 7, wherein the target vector is matched using a perceptual weighting criterion.

9.      A system according to claim 7, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

10.     A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

14

a target vector generator for generating data including a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable amplitude, variable sign, variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulses which, according to a maximum likelihood criterion, most closely represents said target vector.

11.    A system according to claim 10, wherein the target vector is matched using a perceptual weighting criterion; and

wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

12.    A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer connected to an output line of said target vector generator and an output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which, according to the maximum likelihood criterion, most closely represents said target vector.

13.    A system according to claim 12, wherein the target vector is matched using a perceptual weighting criterion.

14.    A system according to claim 13, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

15.    A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer connected to an output line of said target vector generator and an output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which, according to the maximum likelihood criterion, most closely represents said target vector, and

one or more pulse sequence modifiers, each having as input at least a sequence of equal amplitude, variable sign, variably spaced pulses, wherein each said pulse sequence modifier modifies its input sequence and produces as output a sequence of variable amplitude, variable sign, variably spaced pulses.

16.    A system according to claim 15 wherein the pulse sequence modification function is based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

17.    A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector.

18.    A system according to claim 17, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

19.    A system according to claim 18, wherein the target vector is matched using a perceptual weighting criterion.

20.    A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector.

21.    A system according to claim 20, wherein the target vector is matched using a perceptual weighting criterion.

22.    A system according to claim 20, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

23.    A system according to claim 21, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

24.    A system according to claim 21 wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; and characteristics of the input speech signal.

18

25.     A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector, and

one or more pulse-train sequence modifiers, each having as input at least a sequence of equal amplitude, variable sign, variably spaced pulse trains, wherein each said pulse sequence modifier modifies its input sequence and produces as output a sequence of variable amplitude, variable sign, variably spaced pulse trains.


26.     A system according to claim 25, wherein the target vector is matched using a perceptual weighting criterion.


27.     A system according to claim 25, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.


28.     A system according to claim 25, wherein the pulse-train sequence modification function is based on the exponential function.

29.    A system according to claim 25, wherein the pulse-train sequence modification function is based on a linear function.

30.    A system according to claim 25, wherein the pulse-train sequence modification function is based on the short-term characteristics of the input speech signal.

31.    A system according to claim 25, wherein the pulse-train sequence modification is based on the long-term characteristics of the input speech signal.

32.    A system according to claim 25, wherein the pulse-train sequence modification function is based on the excitation signal from previous frames.

## III. Status of Claims

Claims 1-32 are presented for appeal. Claims 1-27 and 29-32 stand rejected under 35 U.S.C. § 103(a) over *Bialik et al.* (U.S. Pat. No. 5,568,588) in view of *Adoul et al.* (U.S. Patent No. 5,754,976); and claim 28 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over *Bialik* in view of *Adoul* and further in view of *Sklar* (Digital Communications Fundamentals and Application). The pending claims under appeal, as presently amended, may be found in the attached Appendix of Appealed Claims.

## IV. Status of Amendments

The application was originally filed on September 8, 1999, including 32 claims. A first Office Action was mailed on January 30, 2002, and in reply, an Office Action Response was filed on July 30, 2002. A final Office Action was mailed on October 22, 2002, and in reply, a Response To Final Office Action and a Notice of Appeal were concurrently filed by facsimile on January 16, 2003. An Advisory Action, which included a new citation (U.S. Patent No. 3,624,302) was mailed on February 20, 2003. On March 17, 2003, an Appeal Brief was filed. An Examiner's Answer was mailed on June 17, 2003, and in reply, a Reply Brief was filed on August 18, 2003. An Office Action was mailed on November 24, 2003, reopening prosecution, and in reply, an Office Action Response was filed on January 14, 2004. A final Office Action was mailed on April 7, 2004, and in reply, an Office Action Response After Final was filed on June 7, 2004. A Notice of Appeal was filed on July 7, 2004, and an Advisory Action was mailed on July 12, 2004.

## V. Summary of Invention

The present invention is directed to a speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector. The method includes generating a plurality of sequences of variable-amplitude pulses from the target vector and the short term characteristics, where each of the sequences has a different average amplitude value, and outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

The present invention is directed to providing a significant improvement on the multi-pulse speech analysis and synthesis ("MPA") teachings of the *Bialik* '588 reference.

Multi-pulse speech analysis and synthesis typically involves dividing the incoming speech signals into frames and then analyzing each frame to determine its representative components, for example, using a frame analyzer to determine the short-term and long-term characteristics of the speech signal. Typically, one, both, or neither of the long- and short-term predictor contributions are subtracted from the input frame, leaving a target vector whose shape has to be characterized from a multiplicity of samples.

As discussed in the background section of Appellant's Specification, one particular MPA approach is described by the '588 reference. The target vector is modeled by a plurality of pulses of equal amplitude, varying location and varying sign (positive and negative). To select each equal-amplitude pulse, a pulse is placed at each sample location and the effect of the pulse, defined by passing the pulse through a filter defined by the LPC coefficients, is determined. The pulse which provides the filter output that most closely matches the target vector is selected and its effect is removed from the target vector, thereby generating a new target vector. The process continues until a predetermined number of pulses have been found. For storage or transmission purposes, the result of the MPA analysis is a collection of pulse locations, pulse signs (positive or negative), and a quantized value of the equal pulse amplitude in each sequence.

In the prior art, the MPA output typically specifies the resulting pulse locations, but not the order in which they were chosen. It also specifies only one gain parameter, so the decoder must reconstruct the pulse sequence using equal amplitudes for all the pulses in the sequence.

According to an example embodiment, the present invention significantly improves over the '588 reference by performing a more accurate MPA analysis; an MPA analysis that, from a maximum-likelihood standpoint, has a much better opportunity for determining the best possible pulse sequence to match the target vector because the pulse sequence is reconstructed using varying amplitude pulses in the sequence. By determining a better match to the target, the perceptual quality of the reconstructed speech is significantly improved.

## VI. Issues for Review

**Issue I:** Are the § 103(a) rejections of claims 1-32 proper when the Examiner failed to present a *prima facie* case of obviousness?

**Issue II:** Is the § 103(a) rejection of claims 1-27 and 29-32 proper when the Examiner's proposed modification of the '588 reference would frustrate the purpose and operation of the '588 reference?

**Issue III:** Is the § 103(a) rejection of claim 28 proper when the Examiner fails to present a *prima facie* rejection by failing to present a combination of references that corresponds to the claimed invention and failing to present evidence of motivation for the modification proposed by the Examiner?

**Issue IV:** Are the § 103(a) rejections of claims 1-32 proper when the Examiner failed to take note of Appellant's arguments presented in the Office Action Response filed on January 14, 2004, and answer the substance thereof, as required by MPEP § 707.07(f)?

## VII. Grouping of Claims

The claims as now presented do not stand and fall together and are separately patentable for the reasons discussed in the Argument. For purposes of this appeal, the claims should be grouped as follows: Group I - claims 1-27 and 29-32; and Group II – claim 28.

## VIII. Argument

Appellant submits that the claims of groups I – II are patentably distinguishable from each other and from the cited prior art references. The claims in group I are patentable over the prior art, because they include subject matter that is not taught or suggested by any of the references cited, including generating from a target vector and short term characteristics, a plurality of sequences of variable-amplitude pulses. The claim of group II is separately patentable over the other claim group because it is directed to subject matter that includes a pulse-train sequence modification function based upon the exponential function, which is not necessarily present in the other claim groups and not taught by the cited prior art.

4

**Issue I:  The § 103(a) rejections of claims 1-32 are not proper when the Examiner failed to present a *prima facie* case of obviousness.**

A.  The § 103(a) Rejections Are Supported By An Unexplained And Illogical Rationale

The Examiner's Section 103(a) rejections fail to present a *prima facie* case of obviousness because Appellant's claims cannot be read to cover the '588 reference, with or without the Examiner's proposed modifications in view of the '976 *Adoul* reference.  With reference to Fig. 1 of the '588 reference, the Examiner alleges that the claim (*e.g.*, claim 1) limitations *"generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses,"* reads on the '588 reference's elements 10, 13, 20 and 38.  Perhaps in view of the clear teaching in each of the '588 embodiments and the '588 claims (*see also* equal-amplitude pulses in Figs. 3A, 3B, 4A and 4B), the Examiner has correctly admitted that none of these cited elements (or any other aspects of the '588 reference) teach *"sequences of variable-amplitude pulses"* (*see* Final Office Action, page 3, last paragraph).

In a hindsight attempt to overcome this deficiency, the Examiner has proposed modifying these cited elements of the '588 reference in view of the alleged teaching of codevector-waveform pulse positions from the '976 reference.  Failing to support the rejection, the Examiner has not, however, explained how this could be accomplished.  Moreover, deduction would suggest that this proposed modification is somehow achieved without modifying elements 10 or 13.  This logic follows since elements 10 and 13 of the '588 reference provide the short-term characteristics (from element 10) and the target vector (from element 13) from which *"a plurality of sequences of variable-amplitude pulses"* are to be generated.  Further, element 20 is described by the '588 reference as merely to "determine[s] the sample location of a first pulse in accordance with [known MPA] multi-pulse analysis techniques" ('588 reference, Col. 4, lines 10-11); as such, it seems untenable that the Examiner would be suggesting that the skilled artisan would be led to change element 20 so that, instead of determining the sample locations of the pulses, it would be modified to generate *"sequences of variable-amplitude pulses."*  Accordingly, the Examiner's proposed modification would have to be achieved by modifying element 38 with the alleged '976 teachings of codevector-waveform pulse positions.

In an abundance of caution, Appellant assumes that the Examiner intended for element 38 to mean element 28, since element 38 is merely an output line that carries the overall result of the analysis (*i.e.*, the pulse sequence that matches the target vector).  With this assumption, the

Examiner's proposed modification would have to be achieved by modifying element 28. Element 28 is the "target vector matcher" that generates the above-mentioned overall result. Thus, Appellant assumes that the Examiner is basing the rejection on the proposed modification of the "target vector matcher 28" with certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference.

Appellant respectfully submits that this proposed modification could not result, even with lengthy research and explanations, in "target vector matcher 28" operating to generate *"sequences of variable-amplitude pulses."* The proposed modification to the "target vector matcher 28" would have to somehow operate as a function of an input signal that changes the amplitude of the pulses in each given sequence. However, with this proposed modification, the "target vector matcher 28" still acts based on the following two inputs: the target vector from element 13, and each pulse sequence provided by line 34. As is typical for every such standard MPA implementation (Col. 1, lines 35-45), the target vector is provided as an input solely as a reference against which the match (estimation) is made, and each pulse sequence provided by line 34 is a sequence of equal amplitude pulses as illustrated, *e.g.*, in Figs. 3A, 3B, 4A, and 4B, the discussions at Col. 2, lines 50-51, Col. 6, lines 8-29, and each issued claim of the '588 reference.

Accordingly, in view of the rationale provided by the Examiner or by deduction, Appellant respectfully submits that this proposed modification does not result in a hypothetical embodiment which corresponds to Appellant's claimed invention including, for example, *"generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses."*

B. The § 103(a) Rationale Would Combine Competing And Incompatible Systems

The Examiner's rejections are based on the MPA type of speech coding approach (as exemplified by the '588 reference) but as somehow modified by certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference. Codevector-waveform pulse positions are used in another type of speech coding system known as the Code-Excited Linear Predictive ("CELP") coding system. Each of the embodiments described by the '976 reference uses and implements the CELP type of coding system as is clearly supported by its "Background of the Invention" section and its detailed description of the algebraic codebook implementations. Indeed, the algebraic codebook implementations correspond to one of the two types of known CELP coding

6

systems described in the "Background of the Invention" section ('976 reference at Col. 2, line 5 *et seq.*).

Appellant respectfully submits that the relied-upon aspects of this '976 CELP coding system are incompatible with the MPA speech coding approach taught by the '588 reference. The differences between these systems are notoriously well known to the skilled artisan, and the Examiner has failed to cite any evidence or even explain what certain aspects relating to the codevector-waveform pulse positions taught by the '976 reference are being used to modify the MPA speech coding approach taught by the '588 reference. Accordingly, these cited references are directed to incompatible speech encoding methods. *See* attached article "Hybrid Codecs".

At page 3 of the final Office Action, the Examiner asserts that the '976 reference teaches an algebraic codebook search method that involves "sequences of variable-amplitude pulses" by citing the encoding principle of the '976 reference. These arguments by the Examiner further show this incompatibility. The Examiner agrees that the purpose of the '976 teachings is to pre-establish a function $S_p$. This is accomplished via the algorithm outlined at Col. 12, line 34 – Col. 14, line 26, which is used to achieve "restraining the subset of codevectors $A_k$ being searched" in the codebook. *See* Col. 14, lines 19-26. However, the '588 reference does not have any such codebook to search. The '588 reference uses target vector matcher 28 on-the-fly and never searches (or mentions) any codebook; this follows as the '588 and '976 methods are entirely different and incompatible aspects of speech coding systems. *See* attached article "Hybrid Codecs" page 2, last paragraph, to page 3, first full paragraph. Further, the Examiner fails to correlate these disparate methods of speech coding. A traditional code-excited linear predictive (CELP) system is shown in the attached block diagram. *See* "CELP Coding of Speech," page 2. The Examiner has failed to identify how the traditional codebooks would be utilized by the '588 method in the proposed combination. Thus, the Examiner has failed to present correspondence to the claimed plurality of sequences of variable-amplitude pulses and failed to present a combination of references that is directed to the same method of speech encoding. Without either of these requirements being met, the Examiner has failed to present a *prima facie* rejection and the Section 103(a) rejections cannot be maintained. Appellant respectfully requests that the rejections be reversed.

> **Issue II:** The § 103(a) rejection of claims 1-27 and 29-32 is not proper when the
> Examiner's proposed modification of the '588 reference would frustrate the purpose
> and operation of the '588 reference.

7

Similar to the previous appeal of these claims, a primary issue is the '588 reference's requirement that each sequence of pulses (or train of pulses) have the same amplitude. Evidence of the significance of this aspect of the '588 reference's teaching, may be seen in that each of the three '588 embodiments as well as each '588 claim includes this requirement. Each of the independent claims is directed to "a plurality of sequences of equal amplitude" (claims 1 and 9), "a sequence of equal amplitudes" (claims 2 and 12), "sign trains of equal amplitude" (claim 5), "variable sign trains of equal amplitude" (claim 10), "trains having the same amplitude level" (claim 15), and "pulses having the same amplitude" (claim 16). With respect to the '588 embodiments, FIGs. 3A, 3B, 4A and 4B, illustrate the equal-amplitude pulse trains and operation of the first embodiment. *See* Col. 2, lines 43-54 (brief description).

In maintaining the prior art rejection, the Examiner improperly attempts to overcome deficiencies in the '588 reference. The Examiner's rejection proposes using the '976 reference's teachings regarding variable amplitude pulses in the '588 reference's processing system.

Appellant has repeatedly shown that the proposed combination of alleged prior art is improper because it would frustrate the purpose and operation of the '588 reference. *See* MPEP § 2143.01 (when a proposed modification would render the teachings being modified unsatisfactory for their intended purpose, then there is no suggestion or motivation to make the proposed modification under 35 U.S.C. § 103(a)).

As with the previous appeal of these claims, the Examiner refuses to accept that the '588 reference requires each sequence of pulses (or train of pulses) to have the same amplitude, as discussed above. The Examiner has not, in any Office Action, explained how the '976 teachings would be combined with the '588 embodiment, thereby failing to comply with 35 U.S.C. § 132 and further, precluding Appellant from considering and responding to the merits of the proposed combination. Notwithstanding this lack of compliance with 35 U.S.C. § 132, Appellant surmised that the proposed modification was to replace the '588 processing of the plurality of single gain pulses with the '976 reference's pulse encoding principle. If the Examiner is suggesting that the differing gain values should be replaced by the '976's amplitude selector 112, which provides a specific function $S_p$ (*see* '976 reference at Col. 12, lines 24-33), the '588's purpose (matching the target vector via performing single gain multi-pulse analysis a number of times) would be destroyed. If the amplitude selector 112 replaces the multiple single gain multi-pulse analysis,

8

then the gain is "pre-established" per the '976 teachings; therefore, the gain is identified without a recurring process (*see* '588 reference Col. 1, lines 49-55 and Col. 2, lines 1-6). This '976 encoding principal would not function at all in the '588 embodiment and, adopting the Office Action's interpretation of the '976 variable-gain encoding principal, the '588 embodiment would no longer have the required single-level pulse sequences. In this regard, the proposed combination would frustrate the operation and purpose of the '588 embodiment. Thus, the Office Action's proposed combination is improper and the rejection cannot be maintained.

Moreover, the Examiner's alleged motivation for the proposed combination is illogical and untenable. The Examiner provides the unsupported statement that the skilled artisan would combine the cited references to obtain a "very good performance" "without paying a heavy price." The Examiner's failure to explain how a "very good performance" would be achieved should be clear in view of the above-discussed resulting inoperable device constructed via this hindsight rejection. The '588 reference teaches an embodiment that attempts to match a target vector by performing single gain multi-pulse analysis a number of times, each with a different gain level. The '976 approach, upon which the Examiner is relying, is an encoding technique that uses a special amplitude selector 112 (Figs. 3A, B and C) to provide a pre-established function (*i.e.*, a pre-established gain) for a pre-assigned relationship to the speech signal (*see* Col. 12, lines 29-33). Replacing the '588 multi-pulse analysis approach (using different gain levels) with the '976 pre-established function would eliminate the recurring process for target vector matching and destroy the '588 method. The Office Action fails to present any evidence of the alleged motivation and the cited teachings would certainly not be motivated. Without a presentation of evidence of motivation, the Section 103(a) rejections cannot be maintained.

> **Issue III:** **The § 103(a) rejection of claim 28 is not proper when the Examiner fails to present a *prima facie* rejection by failing to present a combination of references that corresponds to the claimed invention and failing to present evidence of motivation for the modification proposed by the Examiner.**

The Examiner failed to present a combination of references that correspond to the claimed invention and failed to present evidence of motivation for the proposed modification of the '588 reference. The Examiner further fails to comply with 35 U.S.C. § 132 because no explanation has been given as to how the teachings of *Sklar* would be combined with the above-

9

discussed modified '588 embodiment, and because the Examiner fails to cite any evidence in support of the notion that the skilled artisan would be led by the prior art to implement this asserted combination of teachings. In this regard, the rejection has not afforded Appellant an opportunity to consider and respond to the merits of this proposed combination of three different teachings. *See* 35 U.S.C. § 132. Moreover, the proposed modification (to replace the '588 processing of the plurality of single gain pulses with the '976 reference's pulse encoding principle and also the cited teaching of *Sklar*) would neither correspond to Appellant's claimed invention (as explained above), would frustrate the purpose and teachings of the '588 reference, and would not (contrary to the unexplained assertion in the Office Action) necessarily result in improved output speech quality. Appellant fails to recognize any evidence that has been presented by the Examiner that such a combination of prior art teachings has ever been suggested or even considered.

The Examiner erroneously asserts that the skilled artisan would be lead by the prior art to modify the '588 reference so that it uses an exponential modification function to provide pulses of varying amplitude in each pulse-train sequence because this would allegedly improve output speech quality. Appellant submits that modifying the '588 reference in this regard would not improve output speech quality because the functional blocks described by the '588 reference would still operate under the design principle that the pulses in each pulse-train sequence have the same amplitude. Thus, the Examiner's assertion is illogical.

The Examiner's assertion in this regard would also undermine the operation and objectives of the '588 reference. As stated in the Summary of the '588 reference and discussed above, each pulse sequence has a single gain level and each pulse sequence is processed as this "single gain pulse sequence" (Col. 2, line 11). The Examiner's proposed modification, however, would result in a different set of objectives, in an inaccurate "perceptual weighting filter" (Col. 2, lines 11-12), an inoperable gain selector, and due to a set of unmappable gain levels for each pulse sequence, such pulse sequences which would not be identifiable to "minimize the energy of the error vector and its corresponding gain level" (Col. 2, lines 12-15). Such a destructive combination is improper and fails to indicate the requisite motivation to support a Section 103(a) rejection. The Examiner has not presented a *prima facie* case of rejection; therefore, Appellant submits that the rejection should be reversed.

10

**Issue IV:** The § 103(a) rejections of claims 1-32 are not proper when the Examiner failed to take note of Appellant's arguments presented in the Office Action Response filed on January 14, 2004, and answer the substance thereof, as required by MPEP § 707.07(f).

The Examiner's April 7th Office Action failed to address Appellant's arguments, while repeating the previous rejections. MPEP § 707.07(f) states, in pertinent part, the following:

> Where the requirements are traversed, or suspension thereof requested, the examiner should take proper reference thereto in his or her action on the amendment. Where the applicant traverses any rejection, the examiner should, if he or she repeats the rejection, take note of the applicant's argument and answer the substance of it. If a rejection of record is to be applied to a new or amended claim, specific identification of that ground of rejection, as by citation of the paragraph in the former Office letter in which the rejection was originally stated, should be given.

In this regard, MPEP § 707.07(f) indicates that the Examiner should take note of Appellant's arguments regarding the impropriety of the proposed combination and answer the substance of it. This is consistent with the purpose of aiding the Appellant in judging the propriety of continuing the prosecution, as indicated in 37 C.F.R. § 1.104(a)(2).

In this instance, the Examiner did not comply with this requirement, and Appellant was not afforded the opportunity to judge the propriety of the Section 103(a) rejections and to form a response thereto. For example, the Examiner failed to respond to Appellant's arguments regarding the proposed combination's lack of motivation due to the resulting frustration of the '588 reference's purpose and operation at pages 4-5 in Appellant's January 14th Office Action Response and at pages 2-3 of Appellant's June 7th Office Action Response After Final. Therefore, Appellant requests that the finality of the Office Action mailed on April 7, 2004, be removed, that the Examiner take reference to the Appellant's arguments, and that the Appellant have an opportunity to respond thereto, should the rejection be maintained.

**APPENDIX OF APPEALED CLAIMS (S/N 09/392,124)**

1.　　In a speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, a method of analyzing the input speech signal comprising:

　　　　generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

　　　　outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

2.　　A system according to claim 1, wherein the target vector is matched using a perceptual weighting criterion.

3.　　A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

　　　　means for generating from the target vector and the short term characteristics, a plurality of sequences of variable-amplitude pulses, each of the sequences having a different average amplitude value; and

　　　　means for outputting a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

4.　　 A system according to claim 3, wherein the target vector is matched using a perceptual weighting criterion.

5.　　A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

an analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude pulses, each of said sequences having a different average amplitude value;

the analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude pulses which, according to an error criterion, represents the target vector.

6.      A system according to claim 5, wherein the target vector is matched using a perceptual weighting criterion.

7.      A speech processing system including a signal processor arrangement that analyzes an input speech signal and, in response, generates the short-term characteristics of the input speech signal and a target vector, comprising:

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable-amplitude, variable-sign and variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs;

the multi-pulse analyzer being further adapted to output a signal corresponding to a sequence of equal-amplitude, variable-sign, variably-spaced pulses which, according to a maximum likelihood criterion, most closely represents the target vector.

8.      A system according to claim 7, wherein the target vector is matched using a perceptual weighting criterion.

9.      A system according to claim 7, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

10.     A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

14

a target vector generator for generating data including a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer adapted to receive the target vector and the short term characteristics and to generate a plurality of sequences of variable amplitude, variable sign, variably-spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulses which, according to a maximum likelihood criterion, most closely represents said target vector.

11. A system according to claim 10, wherein the target vector is matched using a perceptual weighting criterion; and

wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

12. A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer connected to an output line of said target vector generator and an output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which, according to the maximum likelihood criterion, most closely represents said target vector.

13. A system according to claim 12, wherein the target vector is matched using a perceptual weighting criterion.

15

14.    A system according to claim 13, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

15.    A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics; and

a multi-pulse analyzer connected to an output line of said target vector generator and an output line of said short term analyzer, wherein said multi-pulse analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulses, each of said sequences having a different average amplitude value, each of said pulses within each sequence having variable amplitudes and variable signs, said multi-pulse analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulses which, according to the maximum likelihood criterion, most closely represents said target vector, and

one or more pulse sequence modifiers, each having as input at least a sequence of equal amplitude, variable sign, variably spaced pulses, wherein each said pulse sequence modifier modifies its input sequence and produces as output a sequence of variable amplitude, variable sign, variably spaced pulses.

16.    A system according to claim 15 wherein the pulse sequence modification function is based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

17. A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of equal amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector.

18. A system according to claim 17, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

19. A system according to claim 18, wherein the target vector is matched using a perceptual weighting criterion.

20. A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector.

21. A system according to claim 20, wherein the target vector is matched using a perceptual weighting criterion.

22. A system according to claim 20, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

23. A system according to claim 21, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

24. A system according to claim 21 wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; and characteristics of the input speech signal.

25.     A speech processing system comprising:

a short-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the short-term characteristics of the input speech signal;

a long-term analyzer that analyzes an input speech signal, and in response to said input speech signal, generates the long-term characteristics of the input speech signal;

a target vector generator for generating a target vector from at least said input speech signal, and optionally, said short-term characteristics, and optionally, said long-term characteristics; and

a pulse-train sequence analyzer connected to at least an output line of said target vector generator and an output line of said short term analyzer, wherein said pulse-train sequence analyzer generates a plurality of sequences of variable amplitude, variable sign, variably spaced pulse trains, each of said sequences having a different average amplitude value, each of said pulse trains within each sequence having variable amplitudes and variable signs, said pulse-train sequence analyzer for outputting a signal corresponding to the sequence of variable amplitude, variable sign, variably spaced pulse trains which, according to the maximum likelihood criterion, most closely represents said target vector, and

one or more pulse-train sequence modifiers, each having as input at least a sequence of equal amplitude, variable sign, variably spaced pulse trains, wherein each said pulse sequence modifier modifies its input sequence and produces as output a sequence of variable amplitude, variable sign, variably spaced pulse trains.

26.     A system according to claim 25, wherein the target vector is matched using a perceptual weighting criterion.

27.     A system according to claim 25, wherein the pulse amplitude variations are based on at least one of: the exponential function; a linear function; the short-term characteristics of the input speech signal; the long-term characteristics of the input speech signal; and the excitation signal from previous frames.

28.     A system according to claim 25, wherein the pulse-train sequence modification function is based on the exponential function.

29.    A system according to claim 25, wherein the pulse-train sequence modification function is based on a linear function.

30.    A system according to claim 25, wherein the pulse-train sequence modification function is based on the short-term characteristics of the input speech signal.

31.    A system according to claim 25, wherein the pulse-train sequence modification is based on the long-term characteristics of the input speech signal.

32.    A system according to claim 25, wherein the pulse-train sequence modification function is based on the excitation signal from previous frames.

# Hybrid Codecs

Hybrid codecs attempt to fill the gap between waveform and source codecs. As described above waveform coders are capable of providing good quality speech at bit rates down to about 16 kbits/s, but are of limited use at rates below this. Vocoders on the other hand can provide intelligible speech at 2.4 kbits/s and below, but cannot provide natural sounding speech at any bit rate. Although other forms of hybrid codecs exist, the most successful and commonly used are time domain Analysis-by-Synthesis (AbS) codecs. Such coders use the same linear prediction filter model of the vocal tract as found in LPC vocoders. However instead of applying a simple two-state, voiced/unvoiced, model to find the necessary input to this filter, the excitation signal is chosen by attempting to match the reconstructed speech waveform as closely as possible to the original speech waveform. AbS codecs were first introduced in 1982 by Atal and Remde with what was to become known as the Multi-Pulse Excited (MPE) codec. Later the Regular-Pulse Excited (RPE), and the Code-Excited Linear Predictive (CELP) codecs were introduced. These coders will be discussed briefly here.

A general model for AbS codecs is shown in Figure 6.
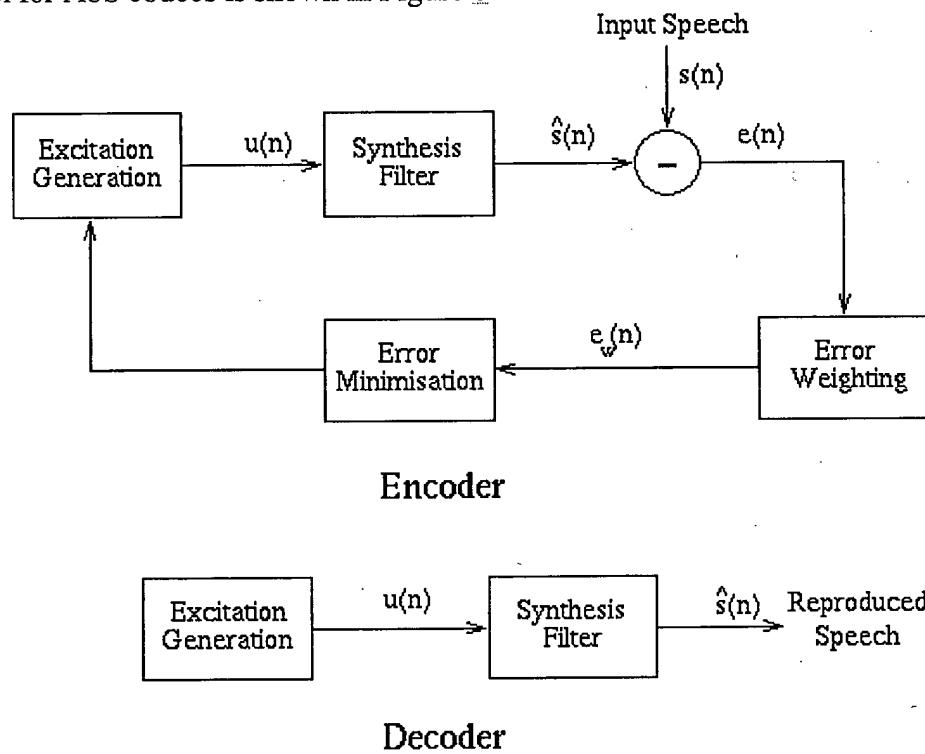


**Figure 6:** AbS Codec Structure

AbS codecs work by splitting the input speech to be coded into frames, typically about 20 ms long. For each frame parameters are determined for a synthesis filter, and then the excitation to this filter is determined. This is done by finding the excitation signal which when passed into the given synthesis filter minimes the error between the input speech and the reconstructed speech. Thus the name Analysis-by-Synthesis - the encoder analyses the input speech by synthesising many different approximations to it. Finally for each frame the encoder transmits information representing the synthesis filter parameters and the excitation to the decoder, and at the decoder the given excitation is passed through the synthesis filter to give the reconstructed speech.

The synthesis filter is usually an all pole, short-term, linear filter of the form

$$H(z) = \frac{1}{A(z)}$$

where

$$A(z) = 1 - \sum_{i=1}^{p} a_i z^{-i}$$

is the prediction error filter determined by minimising the energy of the residual signal produced when the original speech segment is passed through it. The order p of the filter is typically around ten. This filter is intended to model the correlations introduced into the speech by the action of the vocal tract.

The synthesis filter may also include a pitch filter to model the long-term periodicities present in voiced speech. Alternatively these long-term periodicities may be exploited by including an adaptive codebook in the excitation generator so that the excitation signal u(n) includes a component of the form Gu(n-α), where α is the estimated pitch period. Generally MPE and RPE codecs will work without a pitch filter, although their performance will be improved if one is included. For CELP codecs however a pitch filter is extremely important, for reasons discussed below.

The error weighting block is used to shape the spectrum of the error signal in order to reduce the subjective loudness of this error. This is possible because the error signal in frequency regions where the speech has high energy will be at least partially masked by the speech. The weighting filter emphasises the noise in the frequency regions where the speech content is low. Thus minimising the weighted error concentrates the energy of the error signal in frequency regions where the speech has high energy. Therefore the error signal will be at least partially masked by the speech, and so its subjective importance will be reduced. Such weighting is found to produce a significant improvement in the subjective quality of the reconstructed speech for AbS codecs.

The distinguishing feature of AbS codecs is how the excitation waveform u(n) for the synthesis filter is chosen. Conceptually every possible waveform is passed through the filter to see what reconstructed speech signal this excitation would produce. The excitation which gives the minimum weighted error between the original and the reconstructed speech is then chosen by the encoder and used to drive the synthesis filter at the decoder. It is this 'closed-loop' determination of the excitation which allows AbS codecs to produce good quality speech at low bit rates. However the numerical complexity involved in passing every possible excitation signal through the synthesis filter is huge. Usually some means of reducing this complexity, without compromising the performance of the codec too badly, must be found.

The differences between MPE, RPE and CELP codecs arise from the representation of the excitation signal u(n) used. In multi-pulse codecs u(n) is given by a fixed number of non-zero pulses for every frame of speech. The positions of these non-zero pulses within the frame, and their amplitudes, must be determined by the encoder and transmitted to the decoder. In theory it would be possible to find the very best values for all the pulse positions and amplitudes, but this is not practical due to the excessive complexity it would entail. In practice some sub-optimal method of finding the pulse positions and amplitudes must be used. Typically about 4 pulses per 5 ms are used, and this leads to good quality reconstructed speech at a bit-rate of around 10 kbits/s.

Like the MPE codec the Regular Pulse Excited (RPE) codec uses a number of non-zero pulses to give

the excitation signal u(n). However in RPE codecs the pulses are regularly spaced at some fixed interval, and the encoder needs only to determine the position of the first pulse and the amplitude of all the pulses. Therefore less information needs to be transmitted about pulse positions, and so for a given bit rate the RPE codec can use many more non-zero pulses than MPE codecs. For example at a bit rate of about 10 kbits/s around 10 pulses per 5 ms can be used in RPE codecs, compared to 4 pulses for MPE codecs. This allows RPE codecs to give slightly better quality reconstructed speech quality than MPE codecs. However they also tend to be more complex. The pan-European GSM mobile telephone system uses a simplified RPE codec, with long-term prediction, operating at 13 kbits/s to provide toll quality speech.

Although MPE and RPE codecs can provide good quality speech at rates of around 10 kbits/s and higher, they are not suitable for rates much below this. This is due to the large amount of information that must be transmitted about the excitation pulses' positions and amplitudes. If we attempt to reduce the bit rate by using fewer pulses, or coarsely quantizing their amplitudes, the reconstructed speech quality deteriorates rapidly. Currently the most commonly used algorithm for producing good quality speech at rates below 10 kbits/s is Code Excited Linear Prediction (CELP). This approach was proposed by Schroeder and Atal in 1985, and differs from MPE and RPE in that the excitation signal is effectively vector quantized. The excitation is given by an entry from a large vector quantizer codebook, and a gain term to control its power. Typically the codebook index is represented with about 10 bits (to give a codebook size of 1024 entries) and the gain is coded with about 5 bits. Thus the bit rate necessary to transmit the excitation information is greatly reduced - around 15 bits compared to the 47 bits used for example in the GSM RPE codec.

Originally the codebook used in CELP codecs contained white Gaussian sequences. This was because it was assumed that long and short-term predictors would be able to remove nearly all the redundancy from the speech signal to produce a random noise-like residual. Also it was shown that the short-term probability density function (pdf) of this residual was nearly Gaussian. Schroeder and Atal found that using such a codebook to produce the excitation for long and short-term synthesis filters could produce high quality speech. However to choose which codebook entry to use in an analysis-by-synthesis procedure meant that every excitation sequence had to be passed through the synthesis filters to see how close the reconstructed speech it produced would be to the original. This meant the complexity of the original CELP codec was much too high for it to be implemented in real-time - it took 125 seconds of Cray-1 CPU time to process 1 second of the speech signal. Since 1985 much work on reducing the complexity of CELP codecs, mainly through altering the structure of the codebook, has been done. Also large advances have been made with the speed possible from DSP chips, so that now it is relatively easy to implement a real-time CELP codec on a single, low cost, DSP chip. Several important speech coding standards have been defined based on the CELP principle, for example the American Department of Defence (DoD) 4.8 kbits/s codec, and the CCITT low-delay 16 kbits/s codec.

The CELP coding principle has been very successful in producing communications to toll quality speech at bit rates between 4.8 and 16 kbits/s. The CCITT standard 16 kbits/s codec produces speech which is almost indistinguishable from 64 kbits/s log-PCM coded speech, while the DoD 4.8 kbits/s codec gives good communications quality speech. Recently much research has been done on codecs operation below 4.8 kbits/s, with the aim being to produce a codec at 2.4 or 3.6 kbits/s with speech quality equivalent to the 4.8 kbits/s DoD CELP. We will briefly describe here a few of the approaches which seem promising in the search for such a codec.

The CELP codec structure can be improved and used at rates below 4.8 kbits/s by classifying speech segments into one of a number of types (for example voiced, unvoiced and transition frames). The different speech segment types are then coded differently with a specially designed encoder for each type. For example for unvoiced frames the encoder will not use any long-term prediction, whereas for

voiced frames such prediction is vital but the fixed codebook may be less important. Such class-dependent codecs have been shown to be capable of producing reasonable quality speech at rates down to 2.4 kbits/s. Multi-Band Excitation (MBE) codecs work by declaring some regions in the frequency domain as voiced and others as unvoiced. They transmit for each frame a pitch period, spectral magnitude and phase information, and voiced/unvoiced decisions for the harmonics of the fundamental frequency. Originally it was shown that such a structure was capable of producing good quality speech at 8 kbits/s, and since then this rate has been significantly reduced. Finally Kleijn has suggested an approach for coding voiced segments of speech called Prototype Waveform Interpolation (PWI). This works by sending information about a single pitch cycle every 20-30 ms, and using interpolation to reproduce a smoothly varying quasi-periodic waveform for voiced speech segments. Excellent quality reproduced speech can be obtained for voiced speech at rates as low as 3 kbits/s. Such a codec can be combined with a CELP type codec for the unvoiced segments to give good quality speech at rates below 4 kbits/s.

# CELP Coding of Speech

Peter Kabal

McGill University

CELP: Code Excited Linear Prediction

- Example: ITU-T G.723.1 Speech Coder
- Used for teleconferencing, for example by MSN Messenger, Yahoo ...

Good Speech Quality / Low Complexity:

- Original
- Coded at 6.3 kb/s

INTERNATIONAL TELECOMMUNICATION UNION

**ITU-T**
TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.723.1
(03/96)

GENERAL ASPECTS OF DIGITAL
TRANSMISSION SYSTEMS

DUAL RATE SPEECH CODER
FOR MULTIMEDIA COMMUNICATIONS
TRANSMITTING AT 5.3 AND 6.3 Kbit/s

ITU-T Recommendation G.723.1

# CELP Block Diagram



Input Speech

Weighted
Error

Error
Weighting Filter

Reconstructed
Speech

$$\frac{1}{A(z)}$$

LP
Synthesis Filter

Fixed
Codebook

Adaptive
Codebook

KTH

**Speech Signal Processing   2003-11-13**